DOCUMENT RESUME

ED 175 223 EC 120 249

AUTHOR Alvir, Howard P.
TITLE Program Evaluation by Objectives. (You're Held Accountable Only for What You Claim Your Objectives Are!)
PUB DATE Apr 79
NOTE 94p.
AVAILABLE FROM Claire Gelinas, 27 Norwood Street, Albany, New York 12203 (Send self-addressed stamped envelope for recent price of hardcopy)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Educational Accountability; *Evaluation Methods; *Handicapped Children; Individualized Programs; Learning Disabilities; Mentally Handicapped; Models; *Program Evaluation; *Student Evaluation

ABSTRACT

The paper examines ways to choose appropriate objectives for program evaluation in the education of the handicapped. Section I discusses how too much pressure can distort priorities (alignment of short-term objectives with long-term objectives), and lists six steps in determining priorities. Section II applies the principles to learning disabled students through true/false and multiple choice questions. Section III describes a format for factual and practical evaluation of individualized education plans (IEPs). Section IV uses true/false questions to apply principles to the education of mentally subnormal students. The remaining five sections of the document outline components of an accountable, cost-conscious evaluation system. A final section provides examples and explanations of graphic summaries of data. (CL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED175223

PROGRAM EVALUATION BY OBJECTIVES

(You're Held Accountable Only For What You Claim Your Objectives Are!!!)

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

H. Alvir

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TECHNICAL DOCUMENT

© April 25, 1979

HOWARD P. ALVIR, Ph. D.
27 NORWOOD STREET
ALBANY, NEW YORK 12203

EC120249

TABLE OF CONTENTS

	Page
SUMMARY	1
SECTION I	3
PRACTICAL APPLICATIONS	
Pressures vs. Priorities	6
Going Beyond the How-To-Do-It Type of Cookbook Directions	7
Review Quiz	9
Test Questions	11
Answer Key	17
SECTION II	18
LEARNING DIFFICULTIES	
Typical Conclusion	19
True/False Questions	20
Multiple Choice	21
Answer Key	22
Applying All This To The Handicapped	22

Page

SECTION III - - - - - - - - - - - - - - - - - - 23

PUTTING THIS
PROGRAM EVALUATION BY OBJECTIVE
TO WORK FOR YOU

Typical Application - - - - - - - - - - - - - - 24

Communication - - - - - - - - - - - - - - - - - 26

Simplicity - - - - - - - - - - - - - - - - - - 27

A Typical Example - - - - - - - - - - - - - - - 28

An Analysis of the Above Example - - - - - - - - 30

A Few Feed-Back Questions to Measure
Your Comprehension of the Above Example - - - - - - 31

Evaluation Self-Profile - - - - - - - - - - - - - 32

Answer Key - - - - - - - - - - - - - - - - - - 34

Answer Key to Your Own Feed-Back Question - - - - - 35

Specific Evaluation Traits - - - - - - - - - - - - 36

Page

SECTION IV - - - - - - - - - - - - - - - - - - - 38

MENTAL SUB-NORMALITY

Practical Principles - - - - - - - - - - - - - - 39

True/False Questions - - - - - - - - - - - - - - 40

Matching Questions - - - - - - - - - - - - - - - 41

Answer Key - - - - - - - - - - - - - - - - - - - 43

Multiple Choice Questions - - - - - - - - - - - - 44

Mainstreamed Research Designed for the Handicapped - - - - - - - - - - - - - - 45

Answer Key - - - - - - - - - - - - - - - - - - - 46

EARLY RECOGNITION OF LEARNING DISABILITIES - - 47

SECTION V - 48

PUTTING THIS ALL TOGETHER INTO A CONTINUOUS EVALUATION SYSTEM

The Main Implications - - - - - - - - - - - - - - 49

The Next Steps - - - - - - - - - - - - - - - - - 51

A Word of Caution - - - - - - - - - - - - - - - - 52

The Type of Evaluation System We Want - - - - - - - 53

How to Focus Your Energies - - - - - - - - - - - 54

The Advantages of Such a Systematic Approach - - - - - 55

Page

SECTION VI - - - - - - - - - - - - - - - - - - 56

SOME DISCUSSION STARTERS

Thirteen Discussion Starters - - - - - - - - - - 57

Self-Evaluation - - - - - - - - - - - - - - - - 58

SECTION VII - - - - - - - - - - - - - - - - - - 62

ANOTHER EXAMPLE OF
PROGRAM EVALUATION BY OBJECTIVE

Basic Principles - - - - - - - - - - - - - - - 63

Questions and Answers - - - - - - - - - - - - 65

SECTION VIII - - - - - - - - - - - - - - - - - - 69

STILL ANOTHER EXAMPLE OF
PROGRAM EVALUATION BY OBJECTIVE

Basic Facts - - - - - - - - - - - - - - - - - - 70

Questions and Answers - - - - - - - - - - - - 71

Page

SECTION IX - - - - - - - - - - - - - - - - - - - 75

EXAMPLES, DEMONSTRATIONS, AND EXPLANATIONS OF GRAPHIC SUMMARIES OF DATA

Introduction - - - - - - - - - - - - - - - - - 76

Explanation of Ranges - - - - - - - - - - - - - 78

Explanation of Pretest Data - - - - - - - - - - 80

Explanation of Pretests - - - - - - - - - - - - 82

Explanation of Gains - - - - - - - - - - - - - 84

Explanation of Analysis - - - - - - - - - - - - 86

Conclusion - - - - - - - - - - - - - - - - - - 87

SUMMARY

This document applies principles to the education of the handicapped.

The title is Program Evaluation By Objectives. As such, this document shows how to choose appropriate objectives for the education of the handicapped. When the objectives selected are appropriate, program evaluation becomes a systematic process.

In the education of the handicapped, the emphasis is on progress and gains. Thus, the subtitle: You're heid accountable only for what you your objectives are! This means that the successful teacher of the handicapped must choose objective objectives that are both challenging and possible. If the objectives are not challenging, time will be spent achieving nothing. If the objectives are not possible, too much pressure will be developed in chasing after unrealistic objectives.

Section I shows how too much pressure can distort PRIORITIES necessary to educate properly the handicapped.

Section II applies these principles to students whose most observable handicap can be described as LEARNING DIFFICULTIES.

Section III tries to apply these principles by demonstrating an efficient format for factual and practical evaluation of IEP's (Individualized Education Plan).

Summary (cont'd.)

Section IV applies these same principles and practical examples to individuals whose most observable handicap can be described as MENTAL SUBNORMALITY.

Section V recommends an evaluation system that is characterized by both accountability and cost-consciousness.

Section VI gives 13 examples of how to start such an evaluation system.

Section VII underlines six commonly agreed upon principles of evaluation that have particular relevance to the evaluation of programs for the handicapped.

Section VIII gets down to specific facts that are important in developing an evaluation system for the handicapped.

Section IX is written with the lazy reader in mind. Many evaluators want to know exactly what to do immediately. Section IX has been presented in the forms of graphs that document results of the evaluation system described in great detail in the preceding section.

Actually, there is no reason why industrious readers cannot profit from Section IX. Each page in Section IX shows a particular graph that can be used for the evaluation of the handicapped. Each graph is followed by an explanation. In short, Section IX demonstrates and summarizes the results of the program evaluation by objectives approach.

SECTION I

PRACTICAL APPLICATION

This section looks at the following contrasting terms: PRESSURE and PRIORITY. A further distinction is made between two related terms: BRAINSTORMING and BRAIN-POWER. The practical implications of distinguishing between these pairs of terms are explained in detail. To emphasize these distinctions and the practical consequences thereof, a review quiz is supplied.

SECTION I

PRACTICAL APPLICATION

The project has several obvious applications to a handicapped population. Here are some indications of these obvious applications.

IEP stands for individualized education plan. Each of the following portions of Section I have several obvious applications for a handicapped population:

PRESSURES VERSUS PRIORITIES: A PRESSURE is anything that distracts a program for the handicapped from respecting its priority objectives: In short, PRESSURES DISTRACT. On the other hand, a PRIORITY is the alignment of short-term objectives with long-term objectives. The development of a handicapped individual may take a long time, but this is no excuse to become discouraged or to waste time: In short, PRIORITIES MEAN PROGRESS. This is the kind of progress that is evident to the handicapped individual, to the teachers of the handicapped individuals, to the parents and relatives of the handicapped individual, and to the general public who fund the program.

The basic message of BRAINSTORMING VERSUS BRAIN-POWER is simple: BRAINSTORMING means looking at a wide variety of options. BRAIN-POWER means selecting and organizing a few options that seem promising of direct and observable benefits to the handicapped.

Practical Application cont'd.

The six steps presented in this section show how to make an IEP more than a cookbook recipe. These six steps focus on a simple easy to do needs assessment process.

A REVIEW QUIZ at the end of the section enables the reader to look at some practical applications of this needs assessment process.

The answer key at the end of the review quiz provides not only multiple choice answers but short statements. Each short statement is an abbreviated telegraph message. It tries to concentrate in a few words the essence of the most practical applications to a handicapped population.

PROGRAM EVALUATION BY OBJECTIVES

(You're Held Accountable Only for What You Claim Your Objectives Are!!!)

PRESSURES vs PRIORITIES: In planning your program or even a small workshop, there are so many things that you must do that you can easily find yourself neglecting the most important items. The pressure to type up the agenda and the name tags can sometimes make you forget about important priorities of program objectives. The following evaluation method is not telling you to forget about the pressures that must be attended to. This evaluation approach is trying to revitalize your efficiency by concentrating on the highest priorities of any program you undertake.

BRAINSTORMING vs BRAIN-POWER: Many workshop and program planning systems try to resemble a brainstorming session. Very often, they remind one of a situation where somebody absent-mindedly pushed the print button on a computer and the computer printed out everything in its memory. The result was a large pile of paper. This pile of paper was too much for the normal human being to read.

Brainstorming can be transformed into brain-power. Brain-power does more than push the print button on a computer. It sits down and comes up with a computer program. The computer program tells the computer what to do, when to do it, and how to do it. Any time you program a computer, you are really planning in great detail. The result is a viable product. When you apply these principles to your program or workshop, you will have something that is practical and something that makes sense.

GOING BEYOND THE HOW TO DO IT TYPE OF COOKBOOK DIRECTIONS: If you and I walk into a tailor shop to get a tailor-made suit, we have our choice of patterns. This is one method of individualization. Even if we choose the same pattern, both of us must be measured. Our dimensions will help individualize our suits even when they are both made upon the same general pattern. It's the same way with any program. Individual environmental differences must be accounted for in the overall planning. Any attempt to prefabricate a program and ship it elsewhere is doomed if it neglects the individualization process that must precede implementation.

STEP 1: The first step in our approach to program evaluation by objectives is for you to tell us in your own words your goals, all of them. After you've had a chance to express yourself, we want to sum up your goals in very short summary statements. If you feel it necessary, you may wish to indicate those goals which have the highest priority concerned for you personally. At the same time, this would be a good chance for you to tell us why these goals are so important to the success of your project as you envision what you want.

STEP 2: After you've made a separate list of your goals, tell us the evaluation criteria you are currently applying to your program, project, or workshop. Specify the forms you use. Indicate those procedures you find most helpful. Mention also the plans you have for future evaluation. Pinpoint those criteria you feel are used to measured the success of your program. Tell us what you like and don't like about it. The more specific you are, the more helpful this evaluation system can be to you.

STEP 3: List all the resources for your program. Two of the best places to start are those that are specified in your budget and those that are involved in your personnel. Don't forget unique institutional characteristics that furnish the background against which your program must perform.

STEP 4: Look over steps 1, 2, and 3. Make sure you have reduced these lists to clear and measurable components. You're now ready for the integration of this into a working system. Pick out your most important objective. Label this number 1. Look at your list of evaluation criteria. Pick out those criteria which are directly related to your priority objectives. Mark these also number 1. Go next to your list of resources. Identify those which are connected with priority objective number 1 and evaluation criteria for objective number 1. Mark these also with number 1.

STEP 5: Once you have done this, you may find out that you have all three components: objectives, evaluation criteria, and resources allied to the same objectives. If you find out that one of these components is missing, now is the time to start thinking about how you can fill in the missing links.

STEP 6: Repeat this process for goal number 2 and goal number 3. This method of making sure that each goal is backed up with evaluation criteria and adequate resources will help you systematize your program. If you so desire and if you have the time, you may go on to each individual goal of your program. One of the things that will happen after you start setting priorities is discovering that certain of your goals cluster together and become subdivisions of a more important major goal. This is part of the process of establishing program priorities.

REVIEW QUIZ

This Review Quiz is built around 12 principles:

1. A small word added or omitted can make a big difference.

2. Programs must be built upon priorities, not upon pressures.

3. Needless details must be avoided.

4. Selection and organization produce better results than a large pile of data.

5. Tailor-made evaluations are better than ill conceived prefabricated evaluations that copy what someone else did somewhere else for an entirely different purpose.

6. Short goal statements sink in systematically.

7. Common sense evaluation concerns communicate consistency and creativity to community members.

8. Non-budget and non-money resources must be mobilized.

9. An assessment system must tie together objectives, evaluations, and resources.

Review Quiz (2)

cont'd.

10. A systematic approach avoids overlooking an important ingredient.

11. Large goals must be implemented a step at a time.

12. Trivial goals should be avoided. The easy way out is seldom the best solution.

REVIEW QUIZ

(You're Held Accountable for What You Claim Your Objectives Are!!!)

____ 1.01 Look at the above subtitle in parentheses. Compare it with the subtitle of PROGRAM EVALUATION BY OBJECTIVES. The subtitle on this page

A. Does not make sense

B. Is longer

C. Does not contain the word "only"

D. Is much more precise and less ambiguous

____ 1.02 As used herein, the word "pressure" refers to

A. Anything that distracts a program or a workshop from respecting its priority objectives

B. The sum total of all forces used to achieve a ranking of objectives

C. A day-by-day plan which relies upon foresight rather than hindsight

D. Taking care of details that must be done in an orderly and sequential process without hurry or rush

____ 1.03 In this same context, brainstorming is used in a pejorative sense because it is used to symbolize program planning that stresses

A. A well thought out sense of priorities

B. Quantity over quality, all the time forgetting that it is important to cut down on needless details

C. Getting people together to find out what they think about a given project

D. A wide variety of material that is short and easily summarized

____ 1.04 The use of brain-power is compared to the successful use of a computer program. This is done because

A. A large pile of data can be easily printed out

B. A wide variety of reading materials are provided in encyclopedic proportions

C. Such an approach allows one to profit from the unforeseen and unforeseeable benefits of serendipity

D. Your resources are channeled and focused to the solution of the highest ranked priority problem

___ 1.05 Let's take an example. You have attended a conference where you liked the evaluation. You borrowed their forms, questionnaires, and evaulation formats. All you changed was the title from yours to theirs. This will give you an evaluation that is

A. Liable to suffer from the worst defects of prefabrication since its like wearing a suit that was designed for someone else's dimensions

B. A good way to save time and money while at the same time insuring yourself of a successful evaluation procedure

C. Something that can only be done if you are certain that this evaluation procedure worked correctly elsewhere in meeting the unique need of that particular program

D. An inexpensive way to individualize your evaluations to the particular goals you have in mind for your project or workshop

___ 1.06 STEP 1: Your goals statements should be kept short in order to

A. Stress all of the details of exactly what you are getting at

B. Form the basis for a quasi-legal definition

C. Permit your main priority to sink in as the most obvious thing you want

D. Economize on writing and editing time

___ 1.07 STEP 2: Your locally developed evaluation criteria are

A. The forms you and other schools must fill out and send to central headquarters

B. Such things as mean, median, and mode

C. Any type of program data that can be interpreted by chi-square or other correlation techniques

D. The yardsticks that you have developed and used in your own local project or program to determine success or failure

___ 1.08 STEP 3: If the only resources you have listed for your program are those that are included in the budget, you have probably overlooked

A. A number of other resources that could be readily available if only you could obtain more money

B. The unique contribution that can be made to your program by the institution in which it is lodged, its traditions, its prestige, its personnel, and its overall impact on the local community

C. A number of technological improvements that have been made since the budget was written.

D. Very little of any real importance or significant impact on your program

___ 1.09 STEP 4: If you label an objective with number 1, your evaluation criteria with number 1, and your resources with number 1, you are

A. Probably going to get confused because everything will be numbered number 1

B. Stressing to yourselves and others that these objectives, evaluation criteria, and resources are correlated one with the other

C. Getting into a coding system that is much too complicated for the simplicity of the system under description

D. Forgetting about the importance of establishing your hierarchy of priorities

___ 1.10 STEP 5: Many well organized program managers have been astonished in step 4 to discover that they have omitted an important component from one of their high priority objectives. Step 5 helps you

A. To make sure that each high priority objective is backed up with appropriate evaluation criteria and resources

B. Forget about structural weaknesses in your overall system

C. To make up for missing resources for objective 1 by over-developing the resources for objective 2

D. To use one of the evaluation criteria for objective 2 to replace the missing evaluation criteria for objective 1

___ 1.11 STEP 6: It is suggested here that you do which to the following when you find yourself with too many goals, each of which is very minute and detailed?

A. Spell out unique evaluation criteria and individualized resources for each minute goal

B. Try to break down each detailed goal into as many subdivisions as possible

C. Make sure that each individual goal you have is given as much consideration as those goals you rank in your top three priorities

D. Cluster together a number of your related goals so that they become subdivisions of a more important major goal

___ 1.12 Since they are held responsible only for the goals they claim to attempt, certain remiss project managers have chosen very minute and irrelevant goals which are easy to attain. How do you evaluate this?

A. "This is a reasonable and common sense procedure."

B. "This procedure is not the best one to follow, but at least it is an honest awareness of what is happening."

C. "Eventually, this type of project manager will find out that the easy way out is not the best solution to his problem."

D. "This doesn't make much difference in the long run provided the program attains 100% of the goals it claims."

ANSWER KEY

1.01 - C - "Only" and ambiguous

1.02 - A - Pressure distracts

1.03 - B - Cut down needless details

1.04 - D - Brain power is organization and selection

1.05 - A - Tailor-made is better than ill-conceived prefabricated

1.06 - C - STEP 1: Short goal statements sink in

1.07 - D - STEP 2: Evaluation criteria are your yardsticks

1.08 - B - STEP 3: Non-budget resources

1.09 - B - STEP 4: Correlate O, E, and R

1.10 - A - STEP 5: No missing links

1.11 - D - STEP 6: Goal and subgoal

1.12 - C - Don't take the easy way out by choosing trivial goals

SECTION II

LEARNING DIFFICULTIES

The following pages present a condensed summary of available research on learning difficulties. This summary has been placed in the format of a true/false and multiple choice examination. The purpose of these questions is to provide the review of salient scientific findings. These findings have been arranged in a sequential format. The reason for this is to facilitate the transition to the development of an evaluation design.

SECTION II

LEARNING DIFFICULTIES

Experimentation with materials has produced some conclusions that help the teachers of the learning disabled. Here are a few of these conclusions.

1. Low academic performance is not the same as a learning difficulty.
2. If a child fails because an inappropriate teaching method was used, this is not enough to justify the diagnosis of learning difficulty.
3. A learning disorder is not the same thing as a learning disability even though both can be called learning difficulties.
4. SCHOOL PHOBIA is one thing; FEAR OF LEAVING HOME is another.
5. The term learning dysfunction is a highly technical term.
6. Concentrating on needed developmental behavior is more profitable than vainly trying to find out the cause of the learning dysfunction.
7. Operant behavior is more effective than negative punishment and more effective than vague perceptual motor training.

The true/false questions and the multiple choice questions on the following pages help the reader apply all of this to adequate program evaluation of the handicapped.

The answers to the true/false and multiple choice questions are based upon public research. The numbers in parentheses after each question document the official scientific source of this answer. In order to avoid overburdening this text, this documentation is available in another publication. Interested parties may contact the author at the address given on the title page to obtain further documentation.

MAINSTREAMED HANDICAPPED

Learning Difficulties

DIRECTIONS: Answer the following questions.

TRUE-FALSE

1. More than 30% of the school-age population are estimated to perform lower academically than expected.

2. Technically speaking, a child has a learning difficulty if and only if an appropriate teaching method was used and the child failed to learn. However, this definition is very general.

3. A learning disorder is an excessive withdrawal response "away from" learning tasks; this response can generalize to "school phobia."

4. A learning disability is an excessive approach response "to things unrelated to school or learning" which interfere with school-appropriate behavior; this response can generalize to "separation anxiety."

5. It's easier to specify what a learning dysfunction is not than to define what it is.

6. Helpers and teachers of children should give most of their time and efforts to finding out the causes of learning dysfunctions.

7. The evidence in so far seems to indicate that the operant behavior model has helped children with learning dysfunctions more than the perceptual-motor model.

8. Extrinsic reinforcers (such as praise, grades, and smiles) can be effective if they are at first paired with more concrete generalized reinforcers (such as candy, sweets, ice cream, or tokens exchangable for prizes).

9. To succeed against learning dysfunctions, a teacher must ask, "What does the child need to learn?" and then ask, "With which carefully planned small steps should I begin to teach the child?"

MULTIPLE CHOICE

1. How does one accurately measure a child's inherent mental endowment?
   A. Use a standard IQ or intelligence test.
   B. Compare scores on two kinds of achievement tests.
   C. Accept only aptitude test scores that agree with achievement test scores.
   D. Evaluate only classroom performance.
   E. Assume the highest test score available is the child's true potential.
   F. None of the above.

2. Which of the following must be absent to have a valid learning dysfunction diagnosis?
   A. General mental subnormality.
   B. Learning problems.
   C. Visual and hearing impairment.
   D. Unrelated psychological disorders.
   E. Cultural and educational disadvantages.
   F. Under-achievement in academic performance.

3. How do children with learning dysfunctions appear in class?
   A. Uncoordinated
   B. Normal and intelligent
   C. Successful and quiet
   D. Inattentive and easily distracted
   E. Hyperactive and impulsive

4. Which of the following are part of effective remediation of learning dysfunctions?
   A. Detection before the age of 10.
   B. Small step competency-based and success-motivated curricula
   C. Use of all five senses to learn
   D. Systematic reinforcement
   E. Attempts to track down the causes of learning dysfunctions
   F. Specific behavioral goals based upon behavioral assessment
   G. Success with the basic tool subjects

ANSWERS

No. 5

T-F

1. F (97)
2. T (98)
3. T (101-2)
4. T (102-3)
5. T (100-6)
6. F (106)
7. T (110)
8. T (111)
9. T (112)

MC

1. F (98-9)
2. A, C, D, E (104)
3. A, D, E (112,104)
4. B, C, D, F, G (112,106,111)

APPLYING ALL THIS TO THE HANDICAPPED

The above answers are more than multiple choice answers. Each short message can be applied to a handicapped population.

It is recommended that, after reading the chapter and taking the review quiz, each reader discuss how these questions deal with specific principles of needs assessment for the handicapped.

SECTION III

PUTTING THIS

PROGRAM EVALUATION BY OBJECTIVES

TO WORK FOR YOU

The following pages explain application of the evaluation design to a handicapped population. All evaluation of an IEP must be individualized. The important thing is to document progress in such a way that the handicapped individual is aware of specific immediate gains that promise eventual long-term gains. The need is documented for an efficient format, factual honesty, simplicity, and practical applications. Each evaluator is shown how to develop an evaluation self-profile. This profile will save time, energy, and money in documenting worthwhile achievements.

SECTION III

PUTTING THIS

PROGRAM EVALUATION BY OBJECTIVE

TO WORK FOR YOU

Here are some further applications of the project to a handicapped population.

NO COMPARISONS INTENDED WITH OTHERS: The evaluation of an IEP must be individual. Trying to compare the IEP of one individual with another is like comparing apples with oranges. Unique differences are essential.

COMMUNICATION: In addition to paperwork, the IEP requires face to face communication. People who work together can take a very mediocre plan and make a success out of it. After all, once progress has been documented, enough momentum has been generated to improve the development of the handicapped individual.

FORMAT VERSUS FORM: Specific questions are not intended. The main idea is to develop a format that enables each educator to specify objectives based upon needs assessment. After this, the yardsticks against which success or failure will be measured are also specified.

HONESTY VERSUS FLATTERY: If the IEP is not achieved during the first year, this does not mean total failure. The fact is obvious that changes must be made in next year's IEP.

SIMPLICITY: Simplicity means stressing the main priority objective. Simplicity means avoiding needless detours that do not benefit the handicapped individual.

The TYPICAL EXAMPLE has been taken from a non-handicapped environment. This is done to stress one principle: IEP PRACTICES THAT WORK WITH HANDICAPPED INDIVIDUALS CAN BE APPLIED TO NON-HANDICAPPED POPULATIONS.

AN ANALYSIS OF THE ABOVE EXAMPLE: This analysis leads to one conclusion: Several specific objectives myst be chosen. Choosing several specific objectives mean ignoring other objectives that simple are not attempted at the present time. Concentrating on a few objectives allows score to be kept. Such concentration focuses energy on essential points.

EVALUATION SELF-PROFILE: These questions are intended for review. As the answer key makes obvious, there are no right or wrong answers. The answer given for each question allows each educator to develop a map of the individual's evaluation style. Thus, the individual will find out whether the evaluation is centered around the self or the outside environment. The evaluator will determine whether the evaluation is summative or formative. The evaluator will decide exactly how each of these dimensions applies to the IEP in question.

PUTTING THIS PROGRAM EVALUATION BY OBJECTIVES TO WORK FOR YOU

NO COMPARISONS INTENDED WITH OTHERS: This approach will work if you use it to self-evaluate how well you're accomplishing what you claim to accomplish. Don't try to use this evaluation system for comparisons with other programs that may be very much like yours, but not identical. This evaluation is intended to bring out the unique assets of your program.

COMMUNICATION: Some projects are well planned, but they fall on their face when it comes time to carry out these plans. Apparently, the project staff wasn't able to work together. There weren't enough face-to-face encounters. The meetings that were held were too formal to facilitate the type of interpersonal communication necessary to make any organizational chart function. People who work together are able to take a very mediocre plan and make a success out of it. They are able to replan their plans according to newer insight and perceptions of what ought to be done. Projects that have built-in mechanisms for communication are able to make the necessary adaptations to succeed. This type of project, that stresses communication rather than evaluation, is the type of project that doesn't see evaluation as a threat.

FORMAT vs FORM: The evaluation method presented here is trying to stress a format and a method. It wants to show you what you can do to tailor-make your evaluation to the specific objectives of your project. It is not concerned with specifying forms and questionnaires that must be filled out by you in order to compare you with other institutions. In others words, you're held accountable in this system evaluation only for what you claim your objectives are. In more precise terms, you are held

accountable for your objectives as measured by yardsticks agreed upon by you. You specify the objectives. You specify the yardsticks against which you are going to measure success or failure. Even the reporting of your progress is up to you. This informal format is something that can make your program improve everyday you are in operation.

HONESTY vs FLATTERY: Honesty demands that you must be willing to admit errors in your program. You must be flexible enough to bend to the requirements of the program environment. If all your evaluation reveals is flattery, you are missing a tremendous opportunity to share ideas and to improve your program. In order to enter into the spirit of honesty, you must realize that admitting a specific error is saying, "Here is something we can improve." Any time you admit the need for improvement, you are not saying, "The entire program is a flop. It needs to be abandoned." You are giving yourself a firm baseline for future improvement. We all want our programs to look good, but any time our program looks too good, we must ask ourself if we're stressing impartiality or flattery. An honest evaluation pinpoints the next step forward.

SIMPLICITY: Simplicity can be described as stressing your main priority item. Any time you find yourself dispersing your energies among a wide variety of scattered objectives, you are lacking both simplicity and a well balanced sense of priorities. There is only so much you can do in a short time. Take care of the most important items, and you will find that the other aspects of your program tend to take care of themselves provided your top priority item is worth doing. If it's worth doing, it's worth doing well.

A TYPICAL EXAMPLE

This example is meant to illustrate the above evaluation process. In no way, does this example mean to imply that there is only one way to apply this process.

QUESTION: What are you trying to do?

ANSWER: We want to provide the necessary individual counseling to help our students over the difficulties they will encounter in their new college environment.

QUESTION: Why do you want to do this?

ANSWER: We feel that without this counseling, most of our in-coming students in this particular category will get lost. They need special help early in the school year in order to adjust to the demands of college academic life. We don't want their personal and emotional problems getting in the way of their academic out put. We also want to show them the school is interested in more than their academic achievement. We respect them as human beings.

QUESTION: How are you going to determine if you have successfully attained this goal of providing adequate individual counseling?

ANSWER: The assistant director of the program is in charge of making sure that each counselor is in his office at 8:00 A.M. in the morning,ready to counsel. It is his job to make sure that at 4:45 P.M. these guidance counselors are still on the job until 5:00 P.M., the official end of the counseling day. He also has a specific list of lunch hours. Periodically, he makes sure that these lunch hours are enforced. At the end of each week, each counselor submits the names of the students that he has counseled during that week. A total is made of the total number of students counseled.

QUESTION: Are you able to come up with some measurable criterion to make this evaluation more specific?

ANSWER: We feel that each student should be counseled twice a semester in an individual session. We have fifty students in the program. This gives a total of 100 counseling sessions per semester. We feel that there are 15 weeks available each semester for counseling purposes. This means that the full time counselor we have hired must see six or seven students a week to keep on schedule.

QUESTION: Do you make any provision to check up on duplications? Are you certain that if your counselor has conducted 120 individual counseling sessions in that semester that he has not included a large number of duplications? Duplications could mean that several students received many counseling sessions while other students received none.

ANSWER: Gee, I never thought of that before. I think this would be a good idea to include in our criteria. I think the way we will cover that is to send out a follow-up questionnaire to each of the 50 students in the program. We will ask them how often they have had a chance to see an individual counselor.

QUESTION: What resources will you have available to help you in the attainment of these objectives?

ANSWER: We will hire a full time counselor. We will give him an office. I will draw up a weekly reporting form on which to keep track of his counseling activities. I will also draw up a questionnaire to be filled out by students every two months.

AN ANALYSIS OF THE ABOVE EXAMPLE

You will notice that the above example was conducted in a question and answer format. The evaluation model herein proposed does not tell you what to do. It does not lay down a set of unchangeable specifications and guidelines. It asks you questions which force you to reflect. It helps you organize your ideas.

You are held accountable only for what you claim you will accomplish.

The above example could have been prolonged by putting several program directors together with the same objectives. Some of these program directors might have claimed that the stress on office hours is actually detrimental to the counseling necessity to be ready to counsel at all hours of the day or night. Others might feel there is too much emphasis on head counting. Another group of counselors might claim that more stress should be placed on process. Someone might object to the student follow up form as a slightly disguised form of espionage on the professional counselor.

There are many more considerations to the above approach to measuring the effectiveness of this individual counseling component. These questions are good to ask. It is not necessary for the busy director of the local problem to listen to all of these contrary viewpoint. He must choose his own objectives. He must decide for what he's going to be accountable.

Once the local program director has made this decision, the system of evaluation herein proposed helpshim check up on himself. When his goals are not clear and measurable, the system provides him with back-up information to help him do the job. When the director himself is vaguely dissatisfied with his goals and wants something better, the system can tell him what others are doing in similar related fields.

This is the value of the system. It helps you carry out your job better. It makes your job easier. It gives you a systematic format in which you can think through your own problems and solve them the way you want. Such an atmosphere is highly more desirable than that of an inspector who looks over your shoulder watching for the slightest error. Such a system is one simple way of increasing your effectiveness without enlarging your budget.

A FEW FEEDBACK QUESTIONS TO MEASURE YOUR COMPREHENSION OF THE ABOVE EXAMPLE

Look at the following questions numbered from 2.01 to 2.10. Look at the directions. Try to answer each question with one alternative.

Do not try to answer these questions in light of your own philosophy. Try to look at the purpose of the evaluation taken in the previous example. Try to decide how this evaluation would fit into the categories herein mentioned. Do not spend too much time in technicalities. This quesionnaire is placed in the form of a dichotomy or two alternate extremes. The real life situation asserts that most evaluationsare on a continuum. They have shades of both extremes.

Your job in going through these ten questions is trying to decide which pole or extreme is most obvious in the evaluation in question.

EVALUATION SELF-PROFILE

Directions: Look over each of the following questions. Choose the alternative that is most closely related to your style of evaluation. It is persumed that you always use a certain amount of both alternatives. In this self-profile of evaluation, you are to choose the alternative you tend to favor most of the time.

The purpose of this questionnaire is to classify a specific evaluation according to ten evaluation traits.

No attempt is herein made to analyze your underlying philosophy of evaluation.

As you answer each question, you must think of one specific evaluation.

____ 2.01 Who is to evaluate?

A. All the evaluation must be self-evaluation

B. All the evaluation must be done by an outside evaluator

____ 2.02 How do you evaluate?

C. Evaluation must be done everyday

D. Evaluation must be done at the end of a project or a program

____ 2.03 When do you evaluate?

E. Before things happen. I rely mainly upon forethought

F. Once things get going. I rely mainly upon afterthought

____2.04 What do you examine?

G. I stress mainly process evaluation

H. I stress mainly product evaluation

____ 2.05 What criteria do you use to judge?

I. Subjective criteria which give the personal viewpoint

J. Objective criteria which avoid personal bias

____ 2.06 Which question do you ask first?

K. "How well did you do?"

L. "What were your objectives?"

____ 2.07 What kind of evaluation comparisons do you want?

M. "I want to self-evaluate my own program only. I have no intention of comparing my results with those of others."

N. "I am interested in finding out how I am doing in comparison with others on a half dozen or more program variables."

____ 2.08 Which do you stress the most?

O. People and communications

P. Planning and organization

____ 2.09 What kind of evaluation help do you want?

Q. "I want a general method and format that I can adapt to my own specific purposes."

R. "I want to be told exactly which forms and questionnaires are to be filled out to self-evaluate my program."

____ 2.10 How far are you willing to go in admitting errors?

S. "Honesty demands that I pinpoint what's not going right in my program."

T. "A person gets farther if he stresses only the positive elements in any evaluation."

ANSWER KEY

2.01 - A - Self

- B - Outside

2.02 - C - Formative

- D - Summative

2.03 - E - Forethought

- F - Afterthought

2.04 - G - Process

- H - Product

2.05 - I - Subjective

- J - Objective

2.06 - K - Norm

- L - Criterion

2.07 - M - Self-evaluate only

- N - Compare with peers

2.08 - O - People (Communication)

- P - Planning (Organization)

2.09 - Q - Format (Guidelines)

- R - Forms (Questionnaires)

2.10 - S - Admit errors

- T - Stress only the successes

ANSWER KEY TO YOUR OWN FEEDBACK QUESTIONS

The previous page is a general answer key. Since this quesionnaire is intended to classify a specific evaluation according to ten evaluation traits, you will find out that there are no right or wrong answers. This classification scheme is intended to point out salient characteristics of a specific evaluation. The reason for identifying these characteristics is to pool this particular evaluation together with other evaluation that have the same ten characteristics in common. This will increase feedback. It will also allow people to compare their evaluation format with similar techniques and designs developed by people who have eight, nine, or ten traits in common with their own evaluation.

SPECIFIC EVALUATION TRAITS

___ 2.01 The example in question could be done by an outside evaluator. He would be able to check up on the office hours and the number of students per week counseled individually by the counselor. Alternative A answers 2.01.

___ 2.02 This evaluation is done every week. The evaluation is not put off until the end of the program. This indicates that alternative C is the answer to 2.02. Everyday is herein interpreted as meaning after each segment or each week of counseling.

___ 2.03 The evaluation criteria were based upon forethought. Not everything has been considered, but the way the evaluation is set up, there is little room to profit from experience. If the counselor doesn't live up to the criteria, he will be told to put in more time on the job and to counsel more students. Alternative E answers 2.03 for this example.

___ 2.04 The emphasis here is on office hours and head counts. This boils down to product evaluation. Alternative H is the answer to 2.04 for this example.

___ 2.05 Hours spent in the job and the number of individuals counseled refer to objective criteria. Alternative J is the answer to 2.05 for this example.

___ 2.06 Emphasis here is placed upon counseling between six and seven students a week for one full time counselor. No question has been made about whether counselors in other institutions counsel more or less students per week. The stress is on a specific objective or head count. Alternative L is the answer to 2.06 for this example.

____ 2.07 As far as comparisons are concerned, this program stresses self-evaluation in a very limited sense. No comparisons are intended with other institutions. Alternative M is the answer to 2.07 for this example.

____ 2.08 This evaluation would seem to stress people and communications. Look again. The stress is on office hours and head count. Nowhere does the evaluation refer to human communication. Alternative P is the answer to 2.08 for this example.

____ 2.09 Forms and questionnaires are stressed here. Alternative R is the answer to 2.09 in this example.

____ 2.10 The way the evaluation critera are set up, there is no room for error. Alternative T is the answer to 2.10 in this example.

SECTION IV

MENTAL SUBNORMALITY

The following pages give some basic facts about individuals who manifest mental subnormality. These facts have been presented in such a way as to speed up the development of an effective evaluation design. Certain generally accepted principles have been placed in the format of a true/false questionnaire to stimulate discussion. Certain key important professional terms have been placed in the format of a contrast. For each term, a number of multiple choice distinctions have been made. An approach to IQ data has been presented. No claim is made about the thoroughness, reliability, or validity of the IQ. However, the IQ is accepted as a statistical fact. Guidelines have been presented that prevent the misuse of IQ data. Since the greatest tool in preventing the misuse of IQ data is an adequate research design, complete details have been given about the development of control and experimental groups. As an important footnote, specific symptoms have been lifted that help in the early recognition of learning disabled children. These characteristics must be noted in order to avoid misusing and misinterpreting IQ data.

SECTION IV

MENTAL SUB-NORMALITY

The following principles can be applied to teaching handicapped individuals who manifest mental sub-normality.

Adaptive behavior is one of the most convincing documentations of mental sub-normality.

The TRUE/FALSE QUESTIONS on the following pages go over general principles that can be used to provide this documentation. This documentation is essential in establishing the basis for any evaluation of any IEP.

The matching questions are designed to distinguish between contrasting terms. For example, the first question gives several distinctions between arbitrary standards and absolute standards. Similarly, the second question gives obvious distinction between verifiable explanations and unverifiable explanations.

These distinctions between contrasting terms are essential to lay a sound conceptual background for the understanding of program evaluation.

These distinctions make obvious the individualization between different IEP's. Such a distinction avoids the danger of comparing things that should not be compared because they are so essentially different.

To avoid any possible confusion, an answer key has been provided.

The multiple choice questions deal specifically with evaluation details. Stress is laid upon proper evaluation design.

One of the first things to understand in evaluating the handicapped is the concept of IQ. The examples given help document how IQ data are determined. This analysis is important to avoid any statistical contamination of the data.

Even thought the IQ section is only one page, it is recommended that each individual reader spend enough time to acquire complete mastery of the statistics involved.

Another set of multiple choice questions deal with experimental groups and control groups. The examples in each of these experimental and control groups are drawn from the area of the handicapped. These technical distinctions between experimental groups and control groups are essential to the evaluation of IEP's. These technical distinctions must be thoroughly understood to avoid unreliable and invalid evaluation conclusions.

Since young children are difficult to test and since IQ is difficult to measure in individuals with learning disabilities, a separate page has been given to help in the early recognition of learning disabilities. This page explains the unusual patterns of behavior, the lack of coordination, the perception difficulties, the abstract thinking difficulties and other problems. In addition, reference is made to a 16 minute film that helps in the EARLY RECOGNITION OF LEARNING DISABILITIES.

MAINSTREAMED E.M.R.

Mental Subnormality

DIRECTIONS: Answer the following questions.

TRUE-FALSE

1. Mental subnormality is a performance deficit which is associated with impairment in adaptive behavior.

2. Mental subnormality is both a disease entity and a permanent, immutable state.

3. The only reason for subnormal functioning on an intelligence test is stupidity.

4. Familial retardation as explained by Zigler comes from families where many relatives score in the subnormal range of intelligence.

5. Familial retardation is solely hereditary and genetic and is not at all related to environmental impoverishment.

6. Organic retardation refers directly to physiological defect and can occur at all levels of socioeconomic status and measured intelligence.

7. Organic retardation cannot be caused environmentally by such factors as aging, radiation treatment, Rh incompatibility, congenital syphilis, encephalitis, maternal rubella, lead poisoning, or accidental injury to the brain.

8. Any competent special educator can determine at once whether a specific child has familial retardation, organic retardation, or both.

9. The common experience all mentally subnormal children share has to do with their inability to meet the arbitrary expectations of the social environment.

10. Art has a mental age of 10 years and an IQ of 120; Bill has a mental age of 10 years and an IQ of 80. They both have had a fairly well matched history of success and reinforcement since their mental ages are the same.

11. The mentally subnormal can learn.

12. To succeed with subnormal children, one simple rule sums up the secret of progress: "Forget about behavioral assets and deficits and zero in on WHY the child is retarded or on the CAUSE of the mental subnormality."

13. For the mentally subnormal, the only valid IQ test is the electro-encephalo-graph (EEG) test that develops a tracing pattern on a piece of paper.

14. Attempts to establish a list of the retarded child's strengths and weaknesses are a waste of time.

15. An even greater waste of time is any attempt to find out what an individual child likes as a reward for learning.

MATCHING

DIRECTIONS: Some questions have more than one correct match. Find them all. Not all matches are used in all questions.

___ 1. Arbitrary standards
___ 2. Absolute standards
- A. Scale for weight
- B. Ruler for height
- C. Test score for intelligence
- D. Stopwatch for time
- E. Professional judgment for diagnosis
- F. Self-report for description
- G. Parental report for symptoms
- H. Opinions of best friend
- I. Calendar for age

___ 3. Verifiable explanations for incorrect answers on an intelligence exam
___ 4. Unverifiable explanations for incorrect answers on an intelligence exam
- J. The child has never been taught or shown the expected response
- K. The child has never attended school
- L. The child is feebleminded
- M. The child can't learn anything
- N. The child's school is inadequate
- O. The child is inadequate
- P. The child refused to learn
- Q. The child is culturally disadvantaged
- R. The child's intelligence is defective through and through
- S. The child reads slowly
- T. The child understands nothing read aloud or silently
- U. The child was tired, nervous, or sick during the exam
- V. The child takes longer to learn than peers
- W. The child's eyesight is weak
- X. The child is hard of hearing
- Y. The child is either borderline or mildly retarded
- Z. The child is unresponsive

___ 5. Mentally retarded
___ 6. Mentally defective
- A. Those who learn slower than their peers
- B. Fast learners
- C. Those who will never learn anything in the hasty opinion of someone
- D. Those who need slow, patient, and reinforced instruction
- E. Those who are incapable of learning, no matter what one might try, according to an unconfirmed diagnosis
- F. An inaccurate and misleading term when used as synonymous with C and E above
- G. A term that merely summarizes an observation
- H. A term that focuses too much on the child's inner learning capacity, which is almost impossible to measure
- I. A term that focuses on the manner in which the child might learn better

___ 7. Familial retardation
___ 8. Organic retardation
- J. Zigler estimates them to be 75% of all retardates
- K. Found in families where many relatives score low on IQ tests

L. Due to an intellectually impoverished environment
M. Physical defects which severely interfere with intellectual functioning
N. Found mainly in poor families
O. Could be caused by an accident
P. Could be due to a single recessive gene as PKU
Q. Could be prevented by a diet low in phenylalanine
R. All in this group are retarded in exactly the same way
S. Related to a chromosone defect as mongolism or Down's Syndrome
T. Could be caused by radiation, Rh incompatible blood, VD, measles, lead poisoning, or brain injury
U. A unitary disorder
V. A statistically normal phenomenon
W. Caused solely by hereditary factors
X. Traceable in part to genetics in some cases
Y. Traceable in part to environment in some cases
Z. Milder form of retardation

___ 9. Recommended by Ross
___ 10. Not recommended by Ross

A. Blame brain injury first
B. Divide children into two groups: brain-damaged and non-brain-damaged
C. Do not train until the cause of retardation is diagnosed and confirmed
D. Use many tests to confirm whether brain-damage or the environment made a child profoundly retarded
E. Diagnose the strengths and weaknesses of the child's response repertoire
F. Realize that two children with the same mental age (MA) and different IQ's have different reinforcement histories
G. Realize that slow learners often lack confidence in themselves, their successes, and the judgments
H. Realize that slow learners are basically "inner-directed."
I. Admit that some children will never learn, no matter what
J. Insist on an IQ score before starting to teach the child
K. Seek out unique rewards that work with specific individuals
L. Use punishment sparingly since reward is more long lasting
M. Use the right reward for each child
N. Never presume all children like all the same rewards

ANSWER KEY

No. 6 - Mental Subnormality

T-F

1. T (114,133)
2. F (114-7, 128, 133)
3. F (115-6, 133)
4. T (121-2)
5. F (122)
6. T (122-3)
7. F (123)
8. F (124-6, 133)
9. T (113, 117, 126-9, 133)
10. F (127, 133)
11. T (129-33)
12. F (129-33)
13. F (132-3)
14. F (129-33)

MATCH

1. A, B, D, I (113-7)
2. C, E, F, G, H (113-7)
3. J, K, N, Q, S, U, V, W, X (114-6)
4. L, M, O, P, R, T, Y, Z (114-6)
5. A, D, G (116-9)
6. C, D, F, H, I (116-9)
7. J, K, L, N, V, X, Y, Z (121-3)
8. M, O, P, Q, S, T, X, Y (121-3)
9. E, F, G, K, L, M, N
10. A, B, C, D, H, I, J

MULTIPLE CHOICE

Directions: Some questions require more one correct response

1. What is the following Table I?
   - A. A list of intelligence quotients(IQ)
   - B. A list of mental ages (MA)
   - C. A formula to turn MA into IQ
   - D. A table showing how to convert correct answers into MA on one specific test
   - E. The relationship between correct answers into MA on all IQ tests

Table I

Number of questions correct on test 100	The Mental Age (MA) that corresponds to a specific number of questions correct
120	1 year
240	2 years
360	3 years
480	4 years
600	5 years
720	6 years
800	7 years
960	8 years

2. How many questions on test 100 must be answered correctly to get a mental age of one and a half years?
   - A. 60
   - B. 120
   - C. 180
   - D. 240

3. What is the approximate mental age of a child who answers 60 questions correctly?
   - A. Retarded
   - B. Handicapped
   - C. Autistic
   - D. One half year
   - E. Unborn
   - F. One year

4. What is to be done with the following Table II?
   - A. Ignore the missing information
   - B. Guess the missing information
   - C. Compute the missing information after answering quewtions 5-8
   - D. Remember: IQ = 100 (MA) $\div$ IQ

NOTE: IQ = 100 (MA) $\div$ CA

Table II

Name	Number of Questions Correct	Corres. Mental Age(MA)	Chron.l Age (MA)	IQ
Al	480	4	3 yrs	133
Bob	480	A	4	B
Coral	480	C	5	D
Dot	960	E	8	F
Ed	240	G	1½	H
Fran	120	I	1½	J
Gil	360	K	2½	L
Helen	600	M	3½	N
Ian	240	O	4	P
Jane	360	Q	4	R
Karen	600	S	10	T
Larry	720	6	12	50

5. What is Bob's mental age? IQ?

6. What is Gil's MA? IQ?

7. What are the missing values?

8. Which of the following are true?
   - A. Anyone with the same IQ has the same MA?
   - B. Anyone with the same MA has the same IQ
   - C. People with the same IQ are equally smart
   - D. People who answer the same number of questions correctly are given the same mental age

ANSWERS

Mainstreamed IQ Data

Multiple Choice

1. D
2. C
3. D
4. C
5. A=4; B=100
6. K=3; L=120
7. A=4; B=100
   C=4; D=80
   E=8; F=100
   G=2; H=133
   I=1; J=67
   K=3; L=120
   M=5; N=143
   O=2; P=50
   Q=3; R=75
   S=5; T=50
8. D

MAINSTREAMED RESEARCH DESIGN:
EXPERIMENTS WITH:
Experimental Groups
Control Groups

MULTIPLE CHOICE

DIRECTIONS: Some questions require more than one correct response.

1. Which of the following are experimental groups?
   A. The group that got the Crest
   B. The group that didn't get the Crest
   C. The group that got no crest but a toothpaste similar in taste, color, and appearance as a placebo
   D. The handicapped group that was trained with appropriate rewards. Some in this group were in traditional training
   E. The handicapped group that was given traditional training that did not stress rewards and not told they were in an experiment
   F. The handicapped group that was given traditional training that did not stress rewards. All in this group were told they were "part of an experiment "

2. Which of the above are control groups? (Use A through F in question 1.)

3. Which of the above are placebo groups? (Use A through F in question 1.)

4. What are the experimental variables whose effects are being analyzed by examining what happens in groups where it is present and in groups where it is absent?
   A. The effects of toothpaste
   B. The effects of Crest
   C. The effects of brushing teeth
   D. The effects of punishment
   E. The effects of appropriate rewards
   F. The effects of traditional training
   G. The effects of mainstreaming the handicapped
   H. The results of introducing appropriate rewards

5. What is the Hawthorne effect?
   A. The fact that Hawthorne condemned control groups as too expensive
   B. The fact that people improve at least at first when they are told they are part of an experiment
   C. The fact that people seldom change
   D. The fact that innovation, both successful and unsuccessful, often seems to produce spectacular results at first

6. What does it mean when the experimental group and the control group have almost identical effects?
   A. The use of reward was effective
   B. The Hawthorne Effect
   C. The control group did better than the experimental group
   D. The experimental group did better than the control group
   E. The experiment did not find significantly better benefits resulting from the experimental variable
   F. The experimental variable produces significantly better measurable results and improvements
   G. The experimental variable should be condemned and not further tested

7. What does it mean when the experimental group and the placebo group, but not the non-placebo control group, seem to do better at first? (Use A through G in question 6.)

8. What does it mean when the experimental group does better than the placebo control group and the non-placebo control group over a long period of time? Reward was the experimental variable. (Use A through G in question 6.)

9. What does it mean when the experimental group does more poorly than the control group? (Use A through G in question 6.)

10. What must be the only significant difference between the people in the experimental and the control group to have a valid comparison on the effects of the experimental variable?
    A. Age
    B. Mental Age
    C. Health
    D. Motivation
    E. Attitude
    F. The presence of the experimental variable, e.g. appropriate rewards
    G. Handicapping condition
    H. Abnormal condition
    I. IQ or aptitude

11. What must be done when the experimental and control groups vary significantly on important variables that could make the groups incomparable or unsuited for valid comparisons?
    A. Accept the results anyway and mention the discrepancy
    B. Repeat the experiment with comparable groups
    C. Use high powered statistics such as analysis co-variance to correct the error mathematically
    D. Admit the error as soon as it is discovered
    E. Forget about it if it was unintentional

12. What must be done when small or minor differences are found between experimental and control groups?
    (Use A through E in question 11.)

ANSWERS

1. A,D
2. B,C,E,F
3. C,F
4. A,E,H
5. B,D
6. E
7. B
8. A,D,F
9. C,E
10. F
11. B,D
12. A,C,D

SECTION V

PUTTING THIS ALL TOGETHER

INTO

A CONTINUOUS EVALUATION SYSTEM

The following pages focus on the evaluation of educational services destined for handicapped individuals. Tips on how to evaluate the linkages between the services offered and the needs of handicapped individuals are specified. Constant stress is placed upon the need for evaluating self-help rather than a cookbook or partial approach. The point is made that all evaluation systems must have tests, criteria, or yardsticks that spell out exactly what is expected and what is to be evaluated. In such a system, top priorities can be rapidly identified and evaluated.

SECTION V

PUTTING THIS ALL TOGETHER

INTO

A CONTINUOUS EVALUATION SYSTEM

This section summarizes the main implications for program evaluation of educational services destined for handicapped individuals.

THE NEXT STEP: A flexible evaluation system must be established. This system must be based upon the needs of handicapped individuals. This system must meet the criteria set up by professional staff members, by funding agencies, by parents and relatives of the handicapped, and by the self-evident needs of handicapped individuals.

A WORD OF CAUTION: The evaluation recommendations in this publication are specific. However, these recommendations should be viewed as self-help rather than a cookbook approach. This means that individuals facing the daily problems are the ones that should examine the proposed evaluation criteria.

THE TYPE OF EVALUATION SYSTEM WE WANT: The we in this section refers to you the evaluator. This evaluation must have targets that describe exactly what is being evaluated. This evaluation system must have tests or a yardstick that spell out explicit evaluation criteria. This evaluation must have technology or tools and resources to implement the desired assessment.

HOW TO FOCUS YOUR ENERGIES: The evaluator must continually ask two questions: Is this program living up to its objectives? Is this program living up to the objectives of the community in which it is established? Overemphasizing one question at the other is a serious evaluation mistake. If either question is overemphasized or left unanswered, the evaluation energies are not being properly focused.

THE ADVANTAGES OF SUCH A SYSTEMATIC APPROACH: The following benefits must be explicit:

ACCOUNTABILITY.

COST CONSCIOUSNESS.

THE ELIMINATION OF DUPLICATION, AND OF OVERLAPPING AND OUTDATED PROCEDURES.

DAILY RESHAPING OF TARGETS, TESTS, AND TECHNOLOGY.

MEASURABLE OBJECTIVES.

The important element is to determine the top priority. The evaluation must inquire into how well the top priority is being achieved.

THE NEXT STEPS

Now that we have explained our system, we're ready to look closely at your system of self-evaluation. We want you to tell us what you do in this area. Give us your version from your point of view. Give us your reactions and your plans for the future.

Once we've had a chance to explain our system to you and you've had a chance to explain your system to us, we are in a better position to listen to each other creatively. Some of the decisions we may mutually make might be: to buy it, to toss it, to revise it, or to come up with something entirely different. Our options are still open.

It is the authors' feeling that the system herein described is workable, operational and easy to plug into want you're doing. The system applicated here is not a monolithic one. It's flexible enough to fit into any school size. It can co-exist with many different procedural processes.

A WORD OF CAUTION

Don't misinterpret what has been said about this new system. It's really not that new.

We don't want you to start doing new things. We don't want you to increase your work load. We don't want you to start adding man-hours to your already busy week.

We want you to piece together your current tasks into a coherent system. Instead of spending new dollars, this will enable you to establish priorities of what you want to do. This means building on what you've already learned. It means taking another step forward. It's not at all the same thing as going into an entirely different approach.

Look at these following six questions:

1. What do you do?
2. Why do you do it?
3. What resources do you use?
4. What are your goals and objectives?
5. How do you know when you've succeeded?
6. What alternate paths do you have available to you?

THE TYPE OF EVALUATION SYSTEM WE WANT

We want an evaluation system that:

A. Places very few constraints on you.

B. Allows you to set up the evaluation criteria that you feel are appropriate yardsticks for your system.

C. Develops your own local autonomy rather than imposes standardization.

D. Keeps you informed of what your colleagues are doing in similar situations.

E. Provides an administrative tool that gives you the data you need when you want it and in the format you feel is best for decisions you have to make.

This system is composed of three components:

TARGETS: (What is it) (Goals, rationales)

TESTS: (Yardsticks) (Evaluation criteria)

TECHNOLOGY: (Tools) (Resources)

HOW TO FOCUS YOUR ENERGIES

QUESTION: Do most programs tend to under-estimate there own effectiveness?

ANSWER: No. They tend to miss their uniqueness and to misclassify their area of impact.

For example, you claim your preeminence is in one specific area. Outsiders see your progress in other areas that may not be appealing to you because of the image you have of yourself. This means that you are not aware of the full impact you make upon the community and the people who see you in daily life. Your program does have links to and influences on the real world. However, you must be aware of exactly what is your own unique contribution.

THE ADVANTAGES OF SUCH A SYSTEMATIC APPROACH

1. Accountability
2. Cost consciousness
3. Eliminate (a) duplication, (b) overlapping and (c) out-dated procedures
4. Daily reshaping of targets, tests, and technology
5. Measurable objectives

You are not asked to over-tax your limited resources. Do not try to strain doing more than you can. Do not simply add on to your workload You are asked to classify and reassign priorities. This is one effective use of local control.

You are free to come up with your own research design and program self-evaluation. However, central agencies do need details. These details are not subject to prior approval. However, they are subject to both inspection and examination. If you've done a good job, your evaluation can stand close scrutiny.

SECTION VI

SOME DISCUSSION STARTERS

The following pages give thirteen examples of how to liven up a discussion. These discussion starters help get the discussion off to a quick and effective start. By saving time in the beginning, these discussion starters insure the recognition of evaluation ideas that can be implemented. The focus is on meeting the need of any program for the handicapped. The specific need given top priority is the one to be evaluated first.

SECTION VI

SOME DISCUSSION STARTERS

The following four pages give 13 discussion starters. These ideas are intended to simplify the evaluation process. It is not necessary to use each of these 13 discussion starters. However, when the discussion wear down, it is time to try a new starter to give more life and emphasis to the evaluation priorit

1. Evaluation is more effective as a maintenance function than it is as a policing function. It is better to grease the wheels than to spy on people.

2. Accountability means that people want you to show them that you are doing a good job with the trust they have committed to your care.

3. Never misinterpret a difficult question as being critical. People who question you are seriously interested in what you are doing.

4. Acknowledge all partial failures. Everytime you fall down, you must get up.

5. Short-term goals must be evaluated and documented with visable results.

6. The general public is not impressed by complicated statistics. The voter and taxpayer are impressed by results that can be understood and verified by the normal human being.

7. You must do more than claim success for your program. You must document how your program is helping both handicapped and normal human beings.

8. Every human being must realize the unique mix of assets possessed by that person. Then, this person must make use of the best available.

9. Keeping score is one way to get discouraged. However, the point is to avoid being ashamed of partial failure. Partial failure must be corrected the next time around.

10. Experience is a good evaluation tool.

11. Recipients of public dollars must make sure that every penny is well spent.

12. Planning means trying out one alternative among many possibilities. The next step is to make sure that that alternative or a better one is realized.

13. Evaluation is a thermometer; it tells how well the program is doing. Evaluation must also be a thermostat; it must direct the program to even better performance.

SOME DISCUSSION STARTERS for Those Interested in Self-Evaluation

1. Evaluation can be thought of as checking up to make sure that people do the things they claim. This is not always the best approach. It is sometimes better to think of evaluation as greasing the wheels. This means doing what you can to make things flow smoothly and better. Evaluation switches from being a policing function to being a maintenance function.

2. Accountability is a key word nowadays. This is no time to go into scholarly definitions of accountability or its implications. Accountability means that legislators, like most people, want to know how well you are spending their money. They're not asking for behavioral objective or performance objectives or the latest in criterion-referenced testing items. They want you to show them that you are doing a good job with the trust they have committed to your care.

3. People are going to ask a lot of questions about your program. Some of these questions may seem critical, yet you must not interpret these inquiries as threats. Consider each one of these questions an occasion to specify more precisely exactly what you are doing and how well you are doing it.

4. Learning from failures is one of the best ways to develop adequate evaluation criteria. Sometimes a failure will tell you something you can't do. Sometimes a failure will tell you something you should have done differently. Do not consider a partial failure as a symbol of ultimate failure. If you fall down, you must get up.

5. Many of the goals you really want to accomplish can be measured only on a long-term basis. This is not practical for the self-evaluation you must give to your program. You should zero in on the population on campus. These are the people to whom you have greatest access. Find out what they don't like today.

This doesn't mean you must forget about tomorrow in followup studies. It does mean that you must be aware of measurement techniques that can work and that can be implemented immediately without great expense.

6. In its simplest form, an evaluation system can be summed up as "the right questions and the right indices." This means asking yourself, "What are my goals?" It also means asking yourself, "What are my yardsticks?" These two questions will provide you with basic data. They will take your vague goal statements and render them more concrete by specifying what you are going to count, measure, or evaluate. You don't have to rely only on things that can be summed up in numbers. You must rely upon things that can be seen, touched, and verified by the typical man on the street. These people are not impressed by complicated statistics. They are impressed by results that can be understood and verified by the normal human being.

7. After a while, you may want to develop a quantitative justification for the continuation of your program. This does not mean that you must switch over to a strictly numerical statistical technique. You can do justice to the mundane observable items and also to your concerns that can't as easily be measured. There is no need for you to fall into a double recordkeeping system. You don't have to keep one set of records for the inspectors and another set of books for yourself. The accent here is not on double bookkeeping. The accent is upon documentation. We want you to go beyond intuition to stress things that can be seen and looked at by the normal human being.

8. You must go beyond intuitive goals that can be perceived only by yourself through the exercise of your emotional perception. You must use all of the available resources to you. It's like trying to win a friend. You must specify your objectives. Your criteria is rather self-obvious. Then you look at the alternatives available: money, good looks, a fast car, fast talk, a nice

smile, the ability to dance, or the ability to converse interestingly on a variety of topics.

Each one of us has a different mix of assets. We use the best of what we have available.

9. Once you start thinking in terms of measurable terminology, the students will not be blind to the fact that you are starting to count. You must not count in such a way as to be misinterpreted by students. Don't make them feel they're being studied like insects under a microscope or microbes in a test tube. Let them realize that you are looking for visible signs of progress. After a while, they will start coming to you and tell you successes and signs of gains scores that they have made on their own. This is the type of data that will help you come up with an evaluation that gives an accurate picture of your success and of your failure. Failure is not something to be ashamed of; it's something to correct next time around.

10. Anytime we say we need viable and realizable objectives, we are really stressing those performances we want to get from our students as a result of our program. If we can't be specific about these, who can? We must tell exactly what we want. On the other hand we must not get so carried away by objectives that we never do anything. Our planning must be tested in the school of experience. We must find out how well it works; we must find out its limitations.

11. Some people like to receive money from a public agency. They will do anything as long as it's mandated by a dollar sign. This is dangerous. Sometimes this will mean that you do nothing but carry out orders. You will be little more than a puppet. You may forget how unique you are. You may not realize the uniqueness of the contribution you can deliver if you let yourself go.

12. Handling public money is a trust, it's also a very complicated matter. Try to do it in such a way as to stress your creativity and originality. Numbers are not the only measuring tool. There are other evaluation instruments. It's up to you to find them. Concentrate on your goals, then think about your yardsticks. After this you have a wide choice of alternates. Each of these possible combinations of targets, tests, or technology will yield a different product and a different price tag. Each one will give a unique student result. This is where your executive ability comes in. You must choose what you want.

13. Evaluation can become a highly intellectual tool. It can help facilitate the free flow of information. This is good. This may result in books and advanced techniques to be talked about in graduate school. On the other hand there is a practical side to evaluation. It can be used as a thermometer to measure how well the funds are producing. This in turn can affect future commitments of money. Evaluation can also be used as a thermostat. This will apply money to the places where it is needed. If you resort to this to warm a room, don't fall for the thermometer error. Heating a thermometer will raise only the indicated temperature. It won't do very much to increase the warmth in the room; this is where you need a thermostat. It must get the furnaces going. This means that when your evaluation shows the need for progress, you must start doing something beside asking for more money.

SECTION VII

ANOTHER EXAMPLE OF

PROGRAM EVALUATION BY OBJECTIVES

The following pages show what a typical evaluation system can resemble. Focus is placed on six principles that lead to a carefully planned evaluation of programs for the handicapped. For simplicity of comprehension, a question and answer format has been used.

SECTION VII

ANOTHER EXAMPLE OF
PROGRAM EVALUATION BY OBJECTIVES

The following page shows how principles can be applied to the evaluation of programs for the handicapped.

The examples given on the follcwing four pages are intended to show what a typical evaluation system looks like. In order to make these examples more apparent in their application to the handicapped, a few principles are given:

1. The program should limit itself to three top priority objectives. If indeed, these objectives are immediately attained, then three more objectives can be listed as stated. Too many objectives mean too much confusion.

2. There must be some measure of partial success for the handicapped. In some cases, the partial success will take many years. Unless partial success is acknowledged to be success, discouragement can set in.

3. A diploma is not the only measure of success. Partial success could involve preparation for a job or a serious vocational choice.

4. The words drop out and withdrawal should not be used. Better terms should be found to indicate the fact that someone has chosen a partial success area.

5. Committment is an important objective. As such, the abstract definition of committement should be avoided. Committement should be specifically interpreted as the amount of self-help program participants give to eachother in achieving overall objectives.

6. Program success for the handicapped means growth. In now way, should every individual be expected to achieve the same absolute amount of growth. For each individual, a significant step forward should be counted as success.

ANOTHER EXAMPLE OF PROGRAM EVALUATION BY OBJECTIVES

The following example is to show you what can be done. It is not meant to be a perfect example. It's a typical example of how your goals can be made more measurable and more meaningful.

QUESTION: What are the objectives of your program?

ANSWER: Our objectives are three-fold:

First, getting the learners through the program successfully.

Second, developing a sense of committment in each learner.

Third, allowing each learner to become aware of the limits of his ability and to reach out to attain these limits.

QUESTION: Why do you say "getting through the program successfully" instead of "graduating"?

ANSWER: We have many ways to measure the effectiveness of our program. Our goal is not limited to the number of graduates we produce.

QUESTION: Could you spell out some of the criteria you consider when saying a person has gotten through the program successfully?

ANSWER: We have a number of ways of determining success. Here are a few of them:

A. Diplomas

B. Serious vocational choices

C. Serious leisure choices

D. Transfers to community colleges or elsewhere

E. Non-systems successes

F. Going to class, exams, library, and doing assignments

G. Spin-off effects to other participants in the program

H. The willingness to take a risk (mellowing after support)

I. Self-initiated success

J. Campus jobs

K. The number of people we have kept out of prison

L. The number of people we have kept off welfare

M. The number of taxpayers we have put on the tax roles

N. The number of people we have introduced into the mainstream of college life

O. The number of juniors and seniors in our program who can zero-in to help the new population of freshmen and sophomores

QUESTION: I noticed you don't use the words "drop-out" or "withdrawal" from your program. Have you eliminated this concept? Have you also eliminated any possibility of failure?

ANSWER: We have spelled out a long list of successes in order to specify the results of our program. We felt that if we lumped all of these catagories together into drop-outs, we'd be doing a disservice both to the person in the program and to the program itself.

We don't want to brand anyone a failure. We don't want to give the impression to program workers that they have failed in a large number of cases. There are places where we have failed. We keep track of them, we admit our error, and we try to do something about it.

QUESTION: Your second goal was to develop a sense of commitment in the students. What evaluation criteria do you use for this?

ANSWER: When a student demonstrates a sense of commitment, we can recognize it. It's very difficult for us to define it beforehand. Let me give you a couple of examples of what we have recognized as evidence of a true sense of commitment:

A. Students who set up tutorial programs in their former high schools

B. Students who tutor others in the program

C. Students who have made significant gains scores (e.g., going from completely against the program to at least semi-involved)

D. Students who have beneficial spin-offs effects on others in the program

Rather than make this list longer, we feel satisfied with leaving our objective as a sense of commitment. We rely upon the judgment of the administration, the teachers, the staff, and the students themselves in the program to decide whether or not a specific individual has developed a sense of commitment. As more students give us a variety of ways of manifesting this sense of commitment, we will come up with additional specific criteria.

We want to avoid the error of over-specifying this sense of commitment to such an extent that an original contribution will not fit into this evaluation category. We don't want to discourage anyone from discovering his thing and applying it to the development of a worthwhile sense of commitment.

QUESTION: Your third goal was for each individual to become aware of the limits of his own ability and to reach out to them. What evaluation criteria do you use for this?

ANSWER: Again we don't want to come up with necessarily countable results. We want results that mean something significant to the individual concerned. Here are a few examples:

A. We expect the person to have a good conception of what he can do and not do in life. He doesn't necessarily have to talk about it, but he has to build upon this as baseline data.

B. We expect the typical student to go beyond self-exaltation. This means he is expected to realize his limits as well as his potential. We don't want him to conclude falsely that he can succeed in everything.

C. In other words, the student has developed a new concept of what it means to fail. Failure doesn't mean that everything is going against him. It means that a specific situation has been tried and been found wanting. This means the learner is to go elsewhere. If he can't get a degree in a particular program, he is to try other alternatives that are available and open to him.

SECTION VIII

STILL ANOTHER EXAMPLE OF

PROGRAM EVALUATION BY OBJECTIVES

The following pages are based upon eleven facts of evaluation methodology. These facts have been explained in a question and answer format for simplicity. Since the next section will give practical examples, the underlying generalizations have been emphasized in detail in this section.

SECTION VIII

STILL ANOTHER EXAMPLE OF
PROGRAM EVALUATION BY OBJECTIVES

The following example makes obvious certain facts:

1. The first draft of many evaluations are seldom mathematically precise.

2. Rough drafts must be questioned and revised in order to make them actively responsive to the needs of handicapped individuals.

3. As the program progresses, a number of specific yardsticks will emerge that should be used to measure success.

4. It is important that all students have an initial positive experience.

5. This initial positive experience helps students develop a positive self image.

6. In order to succeed, students must learn to be self-critical.

7. Similarly, the program itself should be self-critical. This means acknowledging short-comings as well as success.

8. As it stands, the following four pages provide an example of the evaluation process. However, as a process, this evaluation process should be shortened.

STILL ANOTHER EXAMPLE OF PROGRAM EVALUATION BY OBJECTIVES

This example is intended to show you that it is not necessary to come up with mathematically precise objectives the first time around. You should try to straighten out your thinking. The danger here is setting goals that are too high. However, the advantage is setting up common goals that are set before the individuals involved in the program. This insures that managers of individual components of the program are aware of the higher level goals for which everyone is striving. This makes individual effort an integral part of the common program objectives.

QUESTION: What are the goals of your program?

ANSWER: We want the student to be oral, social, and traditional.

QUESTION: Those sound like good slogans, but what exactly do you mean? Let's start with your meaning of the word "oral."

ANSWER: Our oral goal is to make the student bilingual.

QUESTION: What do you mean by bilingual? Do you want them to speak both French and English? Do you want them to speak both Spanish and Italian? Just what exactly do you mean by bilingual?

ANSWER: In the typical English course on campus, we want the student to recognize the place and function of street language which he already possesses. This is one of the languages he will use. We also want him to learn the place and function of accepted standard English usage.

In this way, he will be able to communicate both with the community from which he springs and with the university community in which he now finds himself. After his time in the university community, he will not find himself returning to his home speaking a foreign language. He will be bilingual. He will be able to communicate his concepts with self-confidence to those people with whom he is in contact daily. He will value their ideas, although he will not necessarily imitate their street language when standard English usage would be more appropriate. He will help those people who can speak only street language to integrate themselves with the larger society.

QUESTION: Have you made your evaluation criteria more specific?

ANSWER: At the present time we are working on it. We think that as the program progresses, we will come up with a number of specific yardsticks that we can use to measure our success in the oral or bilingual goals of the program. At the present time, we think we have a clear idea of what we want. We are waiting for further clarification of the details.

QUESTION: What do you mean by the social goal of your program?

ANSWER: Our social goal stresses community change.

QUESTION: How do you measure the attainment of this goal?

ANSWER: It's not easy, but we have come upon two general guidelines that help us draw up specific evaluation criteria:

A. We want the student to have an initial positive experience. This means he is to experience success in our community. He is to find out that he can go beyond himself and his environment. This will give him the start necessary to become self-motivating.

B. We want him to develop a good self-image. We aren't going to give him a phony type of success. We want him to do something that is difficult, we want him to fall flat on his face. We want him to keep on attacking difficult goals until he can succeed. In this way, success is something he has earned. It is not something given to him on a silver platter. We feel he will appreciate it more this way.

QUESTION: What do you mean by your traditional goal?

ANSWER: By our traditional goal, we are referring to intellectual self-development.

QUESTION: Could you be more specific?

ANSWER: We do feel that there is such a thing as an intellectual life. This refers to a person who knows how to spend his leisure time learning. We have found out that there are going to be at least two principal dimensions in our evaluation of this intellectual life:

A. This life must be self-examined. This means that the learner himself is to start asking the questions of: "What are my goals?"; "What are my evaluation criteria?"; "What are my resources?" He has to start doing this on his own.

B. This life is to be self-critical. It is easy to criticize others. The learner must now zero-in on himself. He must find out his shortcomings and his successes. I could go on and on, but at the present time, we have not come up with more precise evaluation criteria.

QUESTION: Why do you like this particular approach even though it doesn't give you precisely measurable and countable results?

ANSWER: Previously, we came to discussions and then we were told exactly what to do. With this type of self-evaluation by program objectives, we feel that we can start discussing things and learning from our experiences.

We like to keep our goals a little bit general. This gives us enough leeway to improvise and to handle the unforeseeable.

We feel that this stress on local autonomy is just what we need. We now begin to feel that our goals are part of us. They are not something forced on us from outside. It only forces us to be precise. We realize that shaky data will be rejected. We don't have to go to the double recordkeeping system. We know what we want and we keep score the way we think it should be done.

SECTION IX

EXAMPLES, DEMONSTRATIONS, AND EXPLANATIONS OF GRAPHIC SUMMARIES OF DATA

The following pages are assembled two by two. The first page gives an example of a graphic summary of data. The next page gives an explanation of the obvious meanings and interpretations of the preceding page. Each graph has been kept as simple as possible. Each graph is designed to fulfill a specific function. The function is explained on the following page.

INTRODUCTION

The following pages demonstrate and explain how to develop graphical summaries of data such as contact hours, pretest scores, posttest scores, and gain scores.

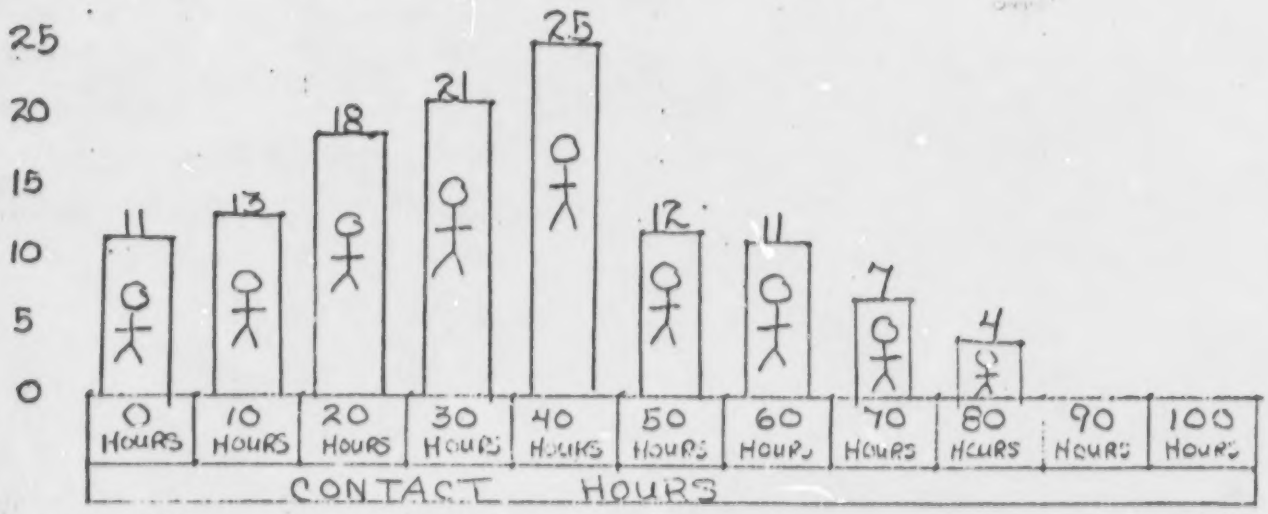
The lefthand page demonstrates what the graph will look like.

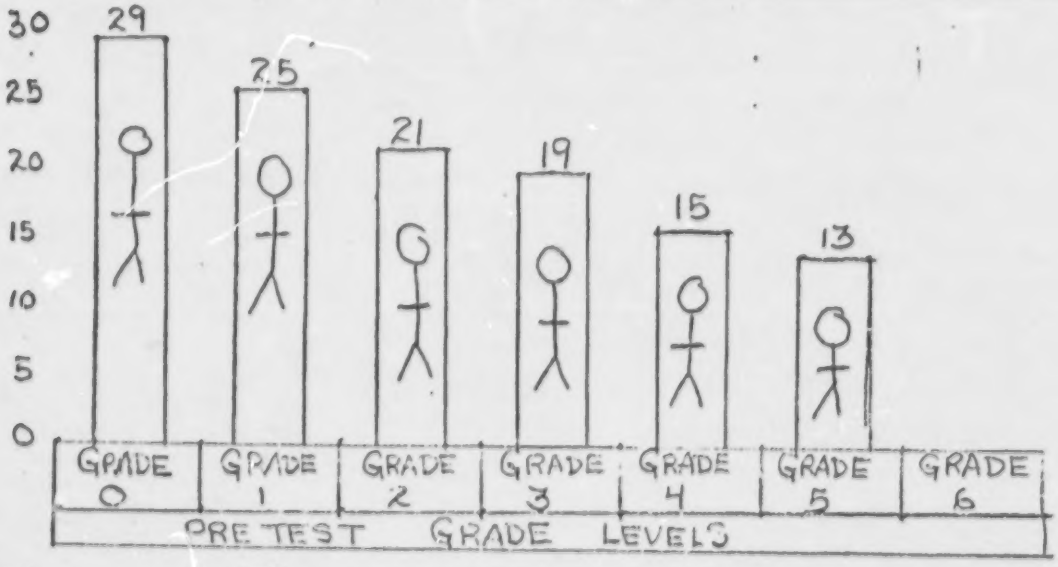
The righthand page explains its meaning and obvious interpretation.

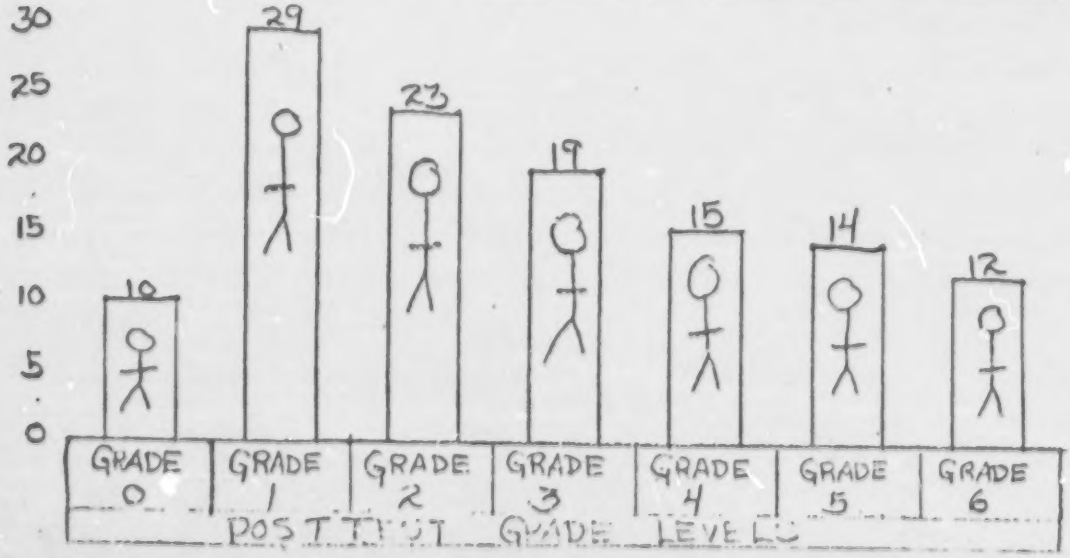
Grade levels have been used for pretests, posttests, and gains. Raw scores could have been used for greater accuracy and for local norming. The danger with grade levels is that of applying the norms of one population to another and different population. Care must be taken to avoid this misuse of norms.

RANGES

... The Number of Individuals in Each Category (TOTAL of 122 People)







EXPLANATION OF

RANGES

THE NUMBER OF INDIVIDUALS IN EACH CATEGORY

(TOTAL OF 122 PEOPLE)

This page illustrates one way to summarize graphically the contact hour, pretest, and posttest data for 122 individuals. The three graphs have been placed on the same page in order to save space and to facilitate comparisons. When a larger number of individuals are analyzed, a separate page may be used for each graph. In this latter case, it still would be advisable to provide a smaller scale graph that contrasts the grade levels for both the pretest and the posttest on the same page.

The important point is to make obvious such pertinent facts that the pretest includes few grade 6 achievers while the posttest contains few grade 0 achievers.

Care must be taken to insure that summary graphical presentations simplify the overview of the data. Clarity of understanding is just as important as arithmetic accuracy. For example, figure drawings inside each column are intended to emphasize that these counts refer to individuals and people. Similarly, the words HOURS and GRADE have been repeated to avoid any possible confusion.

It should be noted that the page described above specifies hours and grade levels without any indication of the subject matter to which this refers. This was done for purposes of illustration. In actual practice, any summary of contact hours and grade levels must be linked to a specific subject matter area or to a specific program.

DATA FOR GRADE TWO PRETEST

Student I.D. Code	CONTACT HOURS	PRETEST GRADE LEVEL	POSTTEST GRADE LEVEL	GAIN SCORE
1	17	1.5	1.6	0.1
2	12	1.5	2.2	0.7
3	15	1.5	2.3	0.8
4	30	1.6	2.5	0.9
5	15	1.7	2.1	0.4
6	18	1.7	2.4	0.7
7	25	1.7	2.2	0.5
8	31	1.8	2.4	0.6
9	43	1.8	2.8	1.0
10	41	1.9	2.8	0.9
11	37	1.9	3.0	1.1
12	31	2.0	3.2	1.2
13	27	2.1	3.2	1.1
14	30	2.1	3.4	1.3
15	33	2.2	3.7	1.5
16	25	2.2	3.9	1.7
17	22	2.2	4.0	1.8
18	31	2.3	3.8	1.5
19	27	2.3	3.6	1.3
20	26	2.3	4.0	1.7
21	21	2.4	4.6	2.2
	AVERAGE = 27.3	AVERAGE = 2.0	AVERAGE = 3.1	AVERAGE = 1.1

EXPLANATION OF
DATA FOR GRADE TWO PRETEST

The purpose of this page is to provide the original data used to analyze the grade two pretest. As is obvious from inspection, the following definitions have been used in this page:

GRADE TWO	refers to all pretest grade levels between 1.5 and 2.4 inclusive.
STUDENT ID CODE	refers to the identification number used in place of the student's name in order to preserve anonymity and confidentiality.
CONTACT HOURS	refer to the number of actual instructional hours each individual student participated in the program. For most students, the first five hours did include initial testing, counseling, diagnosis, prescription, and instructional placement.
PRETEST GRADE LEVEL	refers to the exact score the individual received on the pretest.
POSTTEST GRADE LEVEL	refers to the exact score the individual received on the posttest administered after the number of contact hours indicated for the individual in the program.
GAIN SCORE	refers to the difference between the individual's posttest and pretest.

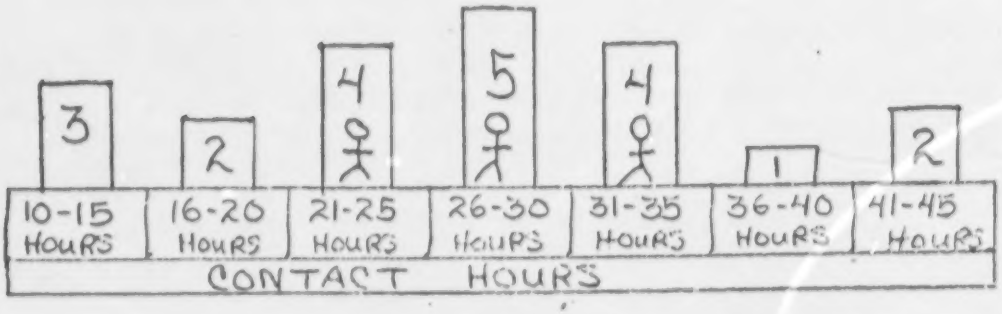
TECHNICAL NOTE

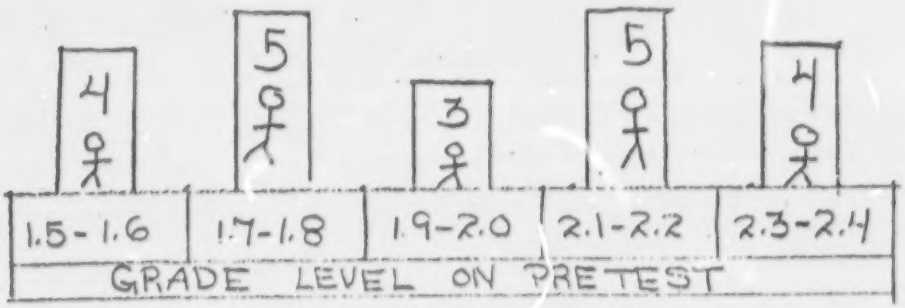
It is customary for evaluators to place the individual data for each student on separate index cards. This permits ranking and sorting on a number of different variables with minimum paperwork and rewriting. As is obvious, on the page explained above, the ranking and sorting was done on the basis of the column entitled PRETEST GRADE LEVEL. This was done in order to facilitate visual correlation between contact hours, pretest grade level, posttest grade level, and gain score.

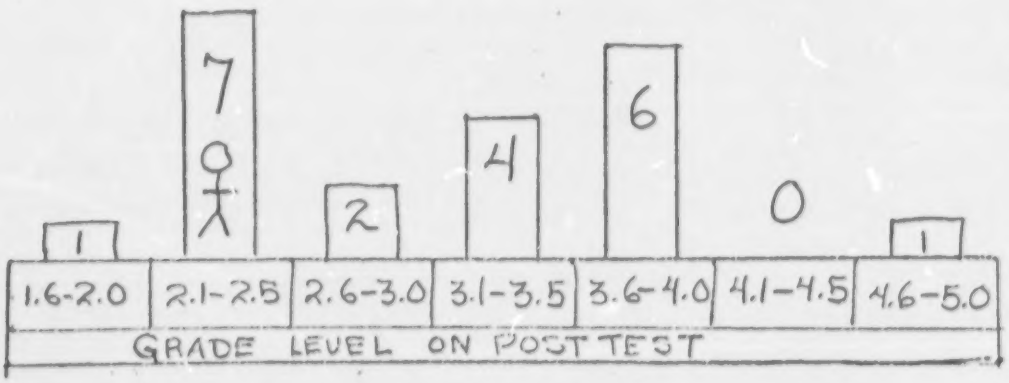
As the following pages will indicate, this correlation can be presented graphically in several different formats, each of which is particularly suited to a specific evaluation purpose.

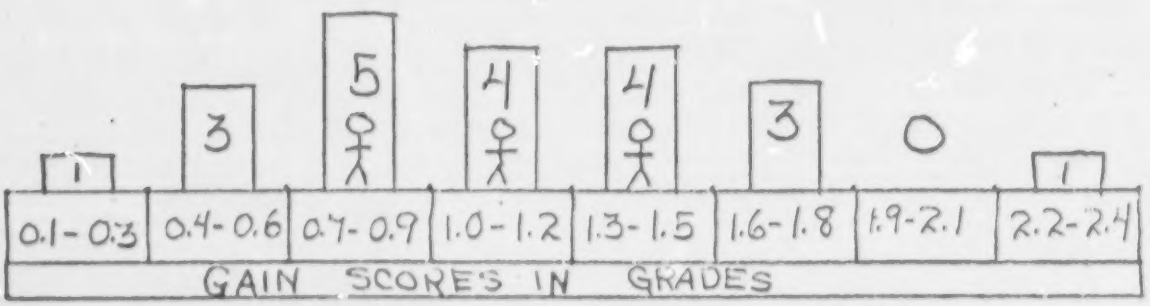
RANGES FOR GRADE TWO PRETEST

The Number of Individuals in Each Category









EXPLANATION OF

RANGES FOR GRADE TWO PRETEST

THE NUMBER OF INDIVIDUALS IN EACH CATEGORY

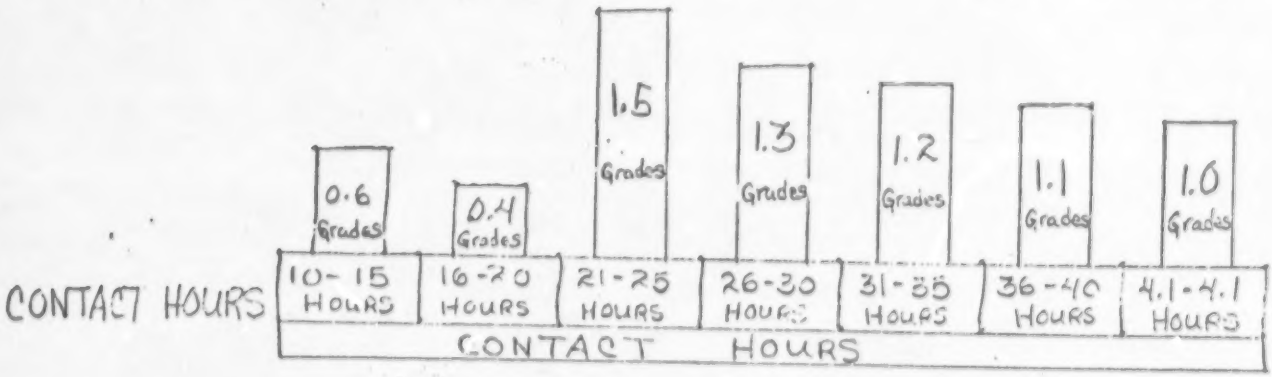
The purpose of this page is to visualize for the grade two pretest the following variables: contact hours, grade level on pretest, grade level on posttest, and gain scores in grades. These items have been previously defined.

There are several reasons for putting all four graphs on the same page. It is easy to compare the four variables for each grade level. It is easy to compare the four grade levels for two or more grades by placing the pages alongside one another. This format forces the evaluator to stress the highlights rather than get lost in statistical detail.

The distribution of each grade level may indicate different intervals for contact hours, grade levels and gain scores. For example, instead of using intervals such as 10-15 and 16-20, certain programs may find it more advisable to use intervals such as 10-19 and 20-29 and so forth. Similarly, instead of 1.5-1.6 and 1.7-1.8 grade levels, other distributions may use different intervals.

GRADE GAIN PER CONTACT HOUR FOR GRADE TWO PRETEST

AVERAGE GRADE GROWTH PER CATEGORY OF CONTACT HOURS



CONCLUSION: The correlation between CONTACT HOURS and GRADE GAIN is low.

SUPPORTING DATA

CONTACT HOURS	PERSONS	ACTUAL GRADE GAIN	AVERAGE
10 - 15	3	0.4 - 0.4 - 0.8	0.6
16 - 20	2	0.1 - 0.7	0.4
21 - 25	4	0.5 - 1.7 - 1.8 - 2.2	1.5
26 - 30	5	0.9 - 1.1 - 1.3 - 1.3 - 1.7	1.3
31 - 35	4	0.6 - 1.2 - 1.5 - 1.5	1.2
36 - 40	1	1.1	1.1
41 - 50	2	0.9 - 1.0	1.0

EXPLANATION OF

GRADE GAIN PER CONTACT HOUR

FOR GRADE TWO PRETEST

AVERAGE GRADE GROWTH PER CATEGORY OF CONTACT HOURS

The purpose of this page is to provide empirical data to answer the following question: "What is the average grade growth for students who spend between 10 and 15 hours in the program, between sixteen and twenty hours in the program, and so forth?"

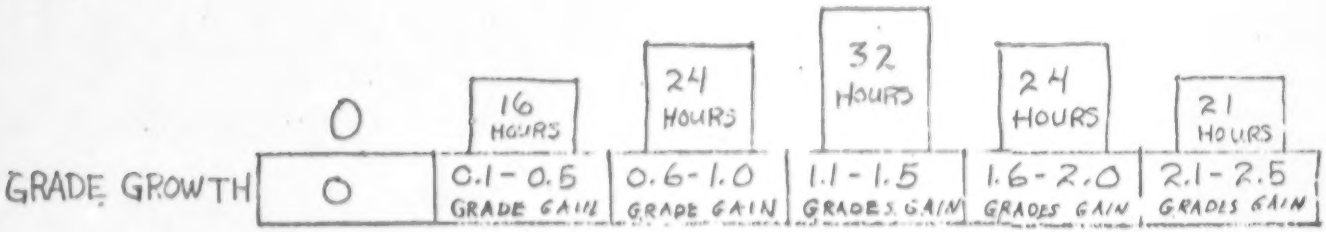
When the correlation between the CONTACT HOURS and GRADE GAIN is high, the grade columns going from left to right will be increasingly higher. This simply means that more contact hours have been found correlated with more grade growth.

In the sample provided on the page described above, it is to be noted that the correlation is low. There are many explanations for this, one of which is the fact that the total number of persons involved in the correlation is 21, a very low number.

The supporting data provided on the bottom of the page is for purposes of simplifying the development of the graphical analysis. In the typical final report, such data are usually assigned to an appendix or an unprinted file for purposes of reference. In the typical final report, such data would be excluded to avoid cluttering up the significant findings. Only the auditors of the evaluation would be specifically interested in the supporting data.

CONTACT HOURS PER GRADE GROWTH FOR GRADE TWO PRETEST

AVERAGE CONTACT HOURS PER CATEGORY OF GRADE GROWTH



CONCLUSION: The correlation between GRADE GROWTH and CONTACT HOURS is low.

SUPPORTING DATA

GRADE GROWTH	PERSONS	ACTUAL CONTACT HOURS	AVERAGE
0	0	0	0
0.1 - 0.5	2	15 - 17	16
0.6 - 1.0	8	12-15-18-25-30-31-41-43	24
1.1 - 1.5	7	27-30-31-31-33-37-37	32
1.6 - 2.0	3	22-25-26	24
2.1 - 2.5	1	21	21

EXPLANATION OF
CONTACT HOURS PER GRADE GROWTH
FOR GRADE TWO PRETEST
AVERAGE CONTACT HOURS PER CATEGORY OF GRADE GROWTH

The purpose of this page is to provide data to answer the following question: "How many contact hours are needed to provide a grade growth of from 0.1 to 0.5 grades, from 0.6 to 1.0 grades, and so forth?"

As is obvious from the graphical data, the correlation between GRADE GROWTH and CONTACT HOURS is low. However, close attention must be given to the first four categories of grade growth. For example, 0 grade growth requires 0 contact hours. Each of the next three categories of grade growth show increasingly large amounts. It is noted that after a grade growth of 1.5, the law of diminishing returns sets in. For all practical purposes this means, that the people who achieved grade growths of more than 1.5 hours did so for reasons other than the total number of contact hours involved. The individuals who made between 1.6 and 2.0 grade growth did so in 24 hours of instruction. The individuals who made between 2.1 and 2.5 grades growth did so in 21 hours.

It is dangerous to interpret graphs based upon small numbers in a very literal fashion. The obvious explanation for the fall-off after 1.5 grade growth is due to the small number of individuals in the last two categories.